Presentation Trainer: What experts and computers can tell about your nonverbal communication

Jan Schneider, Dirk Börner, Peter van Rosmalen, Marcus Specht

Welten Institute, Open University of the Netherlands, Heerlen, The Netherland

Corresponding author

Email: jan.schneider@ou.nl

Address: Valkenburgerweg 177, 6419 AT Heerlen, Netherland

Abstract

The ability to present effectively is essential for professionals; therefore oral communication courses have become part of the curricula for higher education studies. However, speaking in public is still a challenge for many graduates. To tackle this problem, driven by the recent advances in computer vision techniques and prosody analysis, multimodal tools have been designed to support the development of public speaking skills. One of these tools is the *Presentation Trainer*, a research prototype able to provide learners with real-time feedback on a set of nonverbal communication aspects. Despite initial positive evaluations, the application still lacks grounding in a valid assessment model for nonverbal communication aspects in the context of presentations. To come up with such a model, we conducted semi-structured interviews with experts in the public speaking domain. Furthermore, the objective of these interviews was also to have a formative evaluation of the *Presentation Trainer,* analyzing how it suits with common practices for teaching and learning public speaking skills. The results of this study identify 131 nonverbal communication practices that affect the quality of a presentation, and summarize experts' points of view regarding multimodal public speaker instructors.

Introduction

It was February 431 B.C. when Pericles gave his funeral speech and exhorted the people in Athens to live up to the standards set by the deceased (Thucydides). One hundred years later, inspired by Pericles's words, the Greek civilization became one of the most influential in human history. Today, more than 2000 years later, good public speakers still inspire people all over the world. The ability to present effectively is considered to be a core competence for educated professionals (Campbell, Mothersbaugh, Brammer, & Taylor, 2001; Hinton & Kramer, 1998; Parvis, 2001; Smith & Sodano, 2011; Morreale & Pearson, 2008). Policy makers in Europe have recognized this relevance and proposed to all higher education institutions to provide students with presentation skills qualifications (Joint Quality Initiative, 2004).

Research has shown that practice and feedback are fundamental aspects for the development and acquisition of public speaking skills (Van Ginkel, Gulikers, Biemans, & Mulder, 2015). However, opportunities to practice and receive feedback are limited, and graduates often lack the skills to speak in public (Chan, 2011) also due to missing experience and practice. Creating more opportunities to practice and receive the needed feedback through more human assistance is neither feasible nor affordable. Hence, the authors argue for technological solutions to face this problem. Sensor-based environments have become increasingly popular (Swan, 2012) and have shown to support learning through feedback in a great variety of learning scenarios (Schneider, Börner, Van Rosmalen & Specht, 2015a). One of these scenarios is public speaking, where diverse sensor devices, such as depth cameras (Microsoft Kinect) and microphones have been used to develop multimodal research prototypes able to provide learners with feedback regarding their nonverbal communication (Barmaki, & Hughes, 2015; Batrinca,

Stratou, Shapiro, Morency, & Scherer, 2013; Damian, Tan, Baur, Schöning, Luyten, & André, 2015; Dermody & Sutherland, 2015; Schneider, Börner, Van Rosmalen, & Specht, 2015b).

One of these prototypes is the *Presentation Trainer* (PT). The PT supports the training and development of public speaking skills, by presenting the learner with real-time feedback regarding basic nonverbal communication aspects, such as the voice volume, posture, use of pauses and gestures. The study in Schneider et al. (2015b) contains a detailed description of the PT and shows that according to machine-based measurements the PT helped learners to significantly improve their performance. These results show the potential of the PT as a support tool for the development of presentation skills. Nonetheless, the goal of the PT is to ensure supporting learners in delivering better presentations to human audiences, in contrast of improving a machine-based score. Two important missing aspects are preventing the current version of the PT to achieve this goal. The first one is an externally validated model to assess influential nonverbal communication aspects for presentations. The second is a formative evaluation to identify how the use of the PT suits, complements and could enhance current training practices for the development of public speaking skills.

Currently the PT uses a rule-based model to assess the nonverbal communication aspects of presentations. This model is composed of a small set of behaviors that when identified are interpreted as mistakes by the PT. The set of recognized behaviors include: crossing arms, hiding hands, slouching, crossing legs, hopping from one foot to the other, not using enough pauses and gestures, and speaking at an incorrect voice volume. These behaviors can be identified in the vast literature regarding public speaking skills (e.g. Bjerregaard & Compton, 2011; Devito, 2014; Gallo, 2014). However, publications regarding these skills usually lack a formal validation of the ideas and concepts described by the authors. The study in Schreiber, Paul & Shibley (2012)

faced this validation challenge and identified a set of validated rubrics to assess the quality of a presentation. However, the presented assessment regarding the nonverbal communication aspects of a presentation is quite limited and does not identify specific behaviors. It only mentions that the nonverbal communication should align with the message and should avoid being distractive. To contribute to the research of the PT and multimodal public speaker instructors in general, in this study we conducted semi-structured interviews with experts in public speaking. During the interviews we inquired about nonverbal behaviors that affect the quality of a presentation, and in addition, did an expert evaluation of the PT.

Method

Design and Sampling

In this study we conducted semi-structured interviews with 10 experts in public speaking. The group of experts consisted of three females and seven males. Eight of the experts have a Dutch nationality and two of them are British. The age of the experts ranged from 26 to 72 years old. Nine of the experts teach or have taught courses on oral communication skills. Three of them have an acting background; three of them have a personal coaching background; and one of them is a researcher on developing presentation skills.

Instruments and Procedure

We structured the interview in six different phases. The first two phases were designed to introduce the study to the experts and gather their personal information. The third phase of the interview consisted of general questions regarding the nonverbal communication during an oral presentation, such as its relevance and feedback methods used to improve it. The purpose of the fourth phase of the interview was to come up with a set of nonverbal communication behaviors that can be identified as ineffective or good practices during a presentation. The fifth phase of the

interview inquired about the different phases of a presentation and the ineffective and good practices that can be typically identified on each phase. Finally, on the sixth phase of the interview the interviewer showed a live demonstration of the PT and asked the interviewee for impressions and opinions regarding the tool.

Data Collection

The interviews took place in May and June 2015. Nine of the experts were interviewed face-to-face and one of them was interviewed in a videoconference call. One interviewer conducted the 10 interviews, which lasted between 1 and 1.5 hours each. During the interviews an open atmosphere was created where expert and interviewer exchanged information and opinions about the subject. Each interview was audio recorded and then transcribed to a text document.

Data Analysis

To analyze information obtained from the interviews we first organized the transcribed data for each interview according to our own interview guideline, allowing us to individually analyze experts' ideas about: nonverbal communication in general and feedback techniques used to improve those skills, specific nonverbal behaviors that influence the quality of a presentation, particular nonverbal behaviors identified on the different phases of a presentation, and impressions regarding the PT.

We identified the different ideas and concepts from the interview through coding (Rubin & Rubin, 2011) using the NVivo 10[1] software tool. Then we counted the recurrences of the coded ideas and concepts among all the interviews in order to discover commonalities among the different experts.

[1] http://www.qsrinternational.com/product

By analyzing the coded ideas regarding the different phases of the presentation, it was possible to identify that there are some nonverbal behaviors that are unique for these phases and some others that are recurrent for all phases of the presentation. These recurrent behaviors were removed from the particular phases of the presentation and added to a list of nonverbal behaviors in general.

Validity and Reliability

We conducted an external validation (O'Connor & Gibson, 2003) in order to validate our coding process. To conduct this external validation out of the total 284 codes used in the interviews, we randomly selected 20 of these codes together with their corresponding extracted answers given by the experts. We asked eight external reviewers to connect the random codes from the list with the extracted answers from the experts, or suggest a new code in case they did not find a match.

The connections between the codes and the extracted answers conducted by the external reviewers in total had a match of 98% with ours. This high match is a good indicator regarding the reliability and validity of our coding process.

Results

Nonverbal communication ineffective and best practices

Regarding nonverbal communication in general, nine experts claimed it to be very important and one of them considered it as irrelevant. Our study identified four different reasons explaining this relevance. The principal reason supported by seven of the interviewees is that nonverbal communication is the mean to transmit the message. The second identified reason supported by five interviewees is that the nonverbal communication helps the speaker to bond

and create trust with the audience. The third reason, asserted by three experts is that the likeability of the speaker highly depends on her nonverbal communication. Finally, two experts suggested that the nonverbal communication of the speaker supports the content of the presentation.

When experts were asked about how to teach nonverbal communication skills, all experts replied of not being aware of a precise process on how to teach these skills. They replied that the teaching process usually adapts to the particular environment of the learners. Usually presentation skills are taught in a very intensive one-weekend course, or in once a week lessons that last for a whole semester. They can be taught in a group, or in one-on-one coaching sessions. What all experts pointed out is that practice and feedback are necessary to learn these skills.

Regarding the methods used to provide learners with feedback, five of the experts use a technique known as the feedback sandwich technique (Docheff, 1990). In this technique the teacher or peers first name one good aspect about the performance of the student, then an aspect for the student to improve, finishing by stating another good aspect about the student's performance. The main objective of this feedback technique is to help the student to make progress without damaging her self-confidence. One expert does not use the sandwich feedback technique but recommends framing the feedback as positive as possible for similar reasons. Three experts include self-, peer- and teacher-assessment to the their feedback. Pointing out that assessing a presentation is a subjective topic, without a right or wrong way to do it. Therefore, having different feedback sources helps to make the learning experience more comprehensive. One expert uses video recordings as a tool to give feedback. While reviewing the video after the

presentation, students together with teachers can discuss it carefully. Two experts said to have used this feedback technique in the past but stopped using it because it is very time consuming and students usually feel uncomfortable while watching the recordings of their performances. One expert mimics the nonverbal communication of the students, and asks the students to reflect and discuss about it, helping them to become aware of the meaning of their own nonverbal communication.

The analysis of the interviews allowed us to identify 61 nonverbal behaviors that can be interpreted as ineffective communication practices and 70 behaviors that can be interpreted as good practices. These identified nonverbal behaviors can be grouped in seven sets of nonverbal communication aspect: posture, gestures, facial expressions, eye contact, use of stage, voice, and pauses.

Posture

Regarding the posture of the presenter, the most identified ineffective posture practice in the interviews, stated by seven experts, is giving the back to the audience, instead of facing them. Six of the experts mentioned that a common ineffective posture practice is dancing. This dancing behavior communicates to the audience that the presenter is nervous. So the presenter should avoid hopping from one foot to the other. Either from side to side what four of them also called as “Merengue”, or back and forward what four of them called as “Salsa”.

Eight of the experts mentioned the importance of having a posture where the presenter can feel grounded in order to communicate the message with confidence. They mentioned that the feet of the presenter should be between shoulder and waist width firmly on the ground, in

order to become grounded. Three experts commented that it could be ok to move and change posture from time to time, as long as the presenter always returns to this grounded posture after some sentences. Most experts also stated the importance of standing erect in order to display confidence. Keeping the shoulders back and relaxed, the chin up, and the neck back were the behaviors that the experts recommended in order to achieve this erect posture. Most experts also recommended standing with an open posture facing the audience as much as possible in order to transmit that the presenter is communicating with the audience. The list displaying all the identified ineffective and good practices identified for posture is displayed in Appendix A.I.

Gestures

Seven of the experts stated that the biggest problem with gestures during a presentation is not using them. As stated by one of them: "There are no rules for the gestures, they have to be your own, but they have to be there".

Half of the experts mentioned that gestures during a presentation should be bigger than usual face-to-face communication as explained by one of them: "*One has to understand that with gestures and everything, everything on stage should be a bit exaggerated, because it is an abnormal distance for communication. Bigger, slower exaggerated gestures are more useful, and more clear for the audience".* Half of the experts commented that gestures should be used deliberately. They can be used for enumeration e.g. "*When saying first, second, third also use your hands*". Gestures are useful to emphasize or stress important points during the presentation. They help the presenter to paint the picture in the audience mind e.g. "*While mentioning the whole world use big open arms gestures, it gives a physical and mental reflection of what you are doing".* Half of the experts recommended using a gesture and then return to your default or

reset posture for presenting. Four of them reminded that gestures are not universal and that they can be interpreted in many different ways, thus recommended to always vocalize them in order to avoid misinterpretation and confusion. The full list of the identified ineffective and good practices regarding the use of gestures is displayed in Appendix A.II.

Facial Expressions

Considering facial expressions, nine experts stated that presenters should avoid having a blank face throughout the whole presentation. As one of them said: "*You should have an alive facial expression. Smile from time to time even when it is a very serious subject. It is good to see that the presenter is human and not trying desperately to be a professional scientific presenter. Because that is not accessible and the audience loses the attention".* As good practices for facial expressions eight experts said that as a general rule of thumb the presenter should smile from time to time during a presentation. Seven of them gave a warning reminding that the facial expression should be congruent with the content. As one of the experts said: "*You won't smile if you are talking about how the people in South Africa could not get their medicines*". The full list of the identified ineffective and good practices for facial expression is displayed in Appendix A.III.

Eye Contact

Eye contact is another important nonverbal aspect during a presentation. Eight experts identified that one problem that presenters have regarding eye contact is avoiding it. Also eight experts commented about the common ineffective practice of having fixed eye contact with someone in the audience while ignoring the rest.

Ten of the experts commented that a presenter should screen the audience and give as much of eye contact as possible. As one of the experts said: "*Look to your bread, the audience gives you the money, look at them. The trick is to more or less maintain your eye contact a bit behind the center of the audience in the center for a lot of time, but of course keep scanning everybody. And it is ok to directly talk to one person, and then to another."* All the ineffective and good practices regarding eye contact are displayed in Appendix A.IV.

Use of Stage

"*Using the stage with awareness is very powerful and useful, but one needs to know why they are walking around the stage".* Regarding the use of the stage the experts pointed out two ineffective practices. Six of the experts considered standing still behind the computer screen, desk or lectern as a behavior that should be avoided throughout a presentation. Four experts said that moving from one side of the stage to the other without a purpose should also be avoided.

In terms of good practices regarding the use of the stage, four experts noted that the presenter should stand in a place where the audience can see her. Five experts said that moving through the stage with purpose is a very good practice for presenting. One expert recommended to move through the stage according to the particular section of the presentation: "*Support your physical position with the section of the presentation. Move back if you want to create physical distance, when it becomes more theoretical*". Other expert recommended the following: "*Move left and right to communicate time or structure, and back and forward for intensity or intimacy*". The list of identified ineffective and good practices regarding the use of stage is displayed in Appendix A.V.

Voice

Eight experts stated that the biggest problem regarding the use of voice was that so many presenters just talk out-loud instead of speaking to the audience. Eight experts also mentioned that a big problem is to focus only on the content of the presentation and not on how to communicate it to the audience. Half of the experts mentioned that one should avoid filler sounds such as hmms, ahms, and etc. as much as possible, since they are distracting and communicate hesitation.

According to seven of the experts one of the most relevant uses of voice during a presentation is to speak to the audience. As one of the experts mentioned: "*Voice should be projected to the audience, you must speak to them*". The full list of identified ineffective and good practices regarding the use of voice is displayed on Appendix A.VI.

Pauses

"*When people become uncertain on the stage, they have the tendency to go faster, because they think the faster I am the sooner it will be over. They put themselves into a drive and do not pause. It never ever works when you are uncertain slow down, pause*". All experts stated on the interviews that the correct use of pauses is crucial during a presentation.

Six experts recommended pausing for a long period of time after telling something important and after asking any type of question. Half of them stated the importance of having a big pause before introducing a new topic. The full list of identified ineffective and good practices regarding the use of pauses is shown in Appendix A.VII.

Phases of a Presentation

The interviews allowed us to identify six phases in a presentation with particular nonverbal practices. These stages are shown in Table 1.

Table 1. Phases in a Presentation

Phases of a Presentation	**# of experts identifying the phase**
Walking to Stage	4
Settle in Time	10
Introduction	5
Middle	10
Conclusion	8
Questions and Answers	4

The first identified phase of a presentation is walking to the stage. As one expert stated: "*A common mistake while walking to the stage is trying to ignore that the presentation already started".* As good practices for walking to the stage, three experts recommended to walk slow and confident while giving eye contact to the audience.

As a second phase experts identified a settle in time. For this phase all experts agree that one should take their time to settle in before saying the first words. During this phase experts recommend to stand still with both feet firmly on the ground, calm down, take some deep breaths, and then start. All the particularly identified ineffective and good practices for these two phases are displayed in Appendix A.VIII and A.IX.

The following identified phase of the presentation is the introduction. The only particular ineffective communication practice for the introduction stated by one of the experts is starting to

talk with a high pitch. As good for good practices experts explained that this phase has to be very intense, energetic but at the same time in a slow pace. As one of the experts passionately stated: "*If an airplane needs to take off, it needs a take off time. You cannot afford that take off time in a presentation; you have to be flying when you start, and you practice that. You need to have attention with yourself, attention with the audience, and dare to start in a different way. Ask yourself: How can I draw the audience to my story? You need stages of pauses especially at the beginning to draw the audience in; they have no clue what you are going to say. And you do not know where their minds are at the moment. You need to take your time to draw them in. High energy and low pace understanding that it is the first time they hear the story*". The full list of particular identified ineffective and good practices for the introduction of a presentation are displayed in Appendix A.X.

Advancing through a presentation the following phase is the Middle. Seven experts stated that the biggest problem on this phase is that the speech becomes monotonous, as one of the experts stated: "*This is the moment when the autopilot takes over, it becomes monotonous, same cadence all time, I push start and the robot is engaged."*

Regarding the good practices for the middle of a presentation eight of the experts recommended changing dynamics during this phase. Experts gave some examples on how it is possible to change the dynamics of a presentation but explicitly stated that there is not one right way to do it. Some of these examples are to become theatrical for few seconds, move on the stage with purpose, change voice according to the sub-phase of the presentation, etc. One of them suggested to move to the back of the stage and speak very clear and slow when talking

about something theoretical; then come close to the audience and talk at a normal speed when telling an anecdote. The full list of ineffective and good practices for the middle of a presentation is displayed in Appendix A.XI.

The next phase of a presentation is the Conclusion. Six experts said that a common ineffective practice is not ending the conclusion with a full stop, instead the presenter continues speaking and murmuring while waiting for the reaction of the audience. Five experts considered an ineffective practice when the presenter does not signify that the conclusion is coming and it appears too sudden.

Regarding good practices, eight experts suggested taking a couple of breaths and staying quiet for a while before giving the conclusion of the presentation. Six of them stated that the conclusion should be spoken slowly and clearly. Appendix A.XII displays a full list of the particular identified ineffective and good practices for the conclusion.

The final phase identified is Questions & Answers. The most common stated ineffective practice is to focus the attention only on the person asking the question. Therefore four experts recommended to identify and acknowledge the person asking the question, and then give the answer to the whole audience. The particular identified ineffective and good practices for Questions & Answers are displayed in Appendix A.XII

Formative evaluation of the Presentation Trainer

We organized the impressions of the experts regarding the PT in four different categories: Good points, limitations, possible improvements, and practical learning scenarios. The PT

demonstration positively impressed all experts. Figure 1 summarizes in a word-cloud the experts' reactions.

Figure 1. Experts' impressions regarding the *Presentation Trainer*.

Besides the generally positive impressions regarding the *PT*, experts also pointed out some limitations (see Table 2). The biggest limitation stated by all of the experts is that there is not a right way to do a presentation; therefore a machine cannot really assess a presentation. Sometimes a presenter might deliberately break a rule and that does not mean that the performance went wrong. As some expert said: "*There is a risk of interventions not always making sense*". Another expert commented: "*Every person is different and what works for someone might not work for the other. Without the teacher I found it very difficult. In general you can't put rules*." Following this line of argumentation, seven experts stated that the PT cannot be used as a substitution for a human tutor. Moreover, half of the experts remarked that

nonverbal communication is tightly connected with content. Without understanding the content is impossible to make a right assessment about the nonverbal communication.

Table 2. Limitations regarding the PT according to experts.

Limitations	# Of Experts comments
No right way to do presentations	10
It cannot substitute a human tutor	7
No connection with content	5
Important to have real public	2
Kinect is not so accessible	2

During the interviews experts were keen on suggesting improvements for the PT. The list of the suggested improvements is displayed on Table 3. Since there is not a right way to do a presentation, eight of the experts suggested that the Presentation Trainer should shift focus and become a tool to develop awareness of nonverbal communication, instead of correcting it. To support this development of awareness, experts suggested improving the PT with the capacity to ask questions, which allow the user to reflect about her performance. One expert said: "*You could use it as if it was curious audience asking you why you did certain things, instead of a perfect instructor"*. Continuing with the paradigm of creating a tool to raise awareness rather than a tool to correct behavior, three of the experts suggested switching the interventions from corrections to warnings letting the users decide whether their behavior was correct or wrong. Two experts proposed the PT to have configurable feedback rules where the teacher or user can set the type of behaviors that should be displayed and avoided for the specific type of presentation.

Four experts commented about adding a timeline at the end showing an overview of the presentation. One of them suggested that this overview could be sent to the teacher, helping the teacher to know what type of exercises and feedback to give to the student in the following lessons. Four of them also commented about the inclusion of videos showing how certain behaviors could be displayed during a presentation.

Table 3. Improvements for the PT according to experts.

Improvements	# Of Experts comments
Develop awareness	8
Presentation Trainer asking questions	5
Timeline at the end	4
Inclusion of training videos	4
Warnings instead of corrections	3
Exercises to practice one skill at the time	2
Configuration of the feedback rules	2
Configuration of the frequency of feedback	1
Patent	1
Levels of difficulty	1

Discussion

General nonverbal communication behaviors

The opinion of the experts regarding the relevance of the nonverbal communication for public speaking aligns with the information found in previous studies (Quianthy & Hefferin, 1999; Van Ginkel et al., 2015), stating that the nonverbal communication is a very important aspect in presentations. More important, however, the interviews with the experts allowed us to

identify a substantial set of nonverbal behaviors that affect the quality of a presentation, making it possible to separate these behaviors into ineffective and good practices. It is important to note that while asking for these behaviors most experts continuously remarked that what they told is based on personal opinions and that one should not take these opinions as laws, since all nonverbal behaviors can be considered correct as long as they align with the message that the presenter wants to transmit. Though, in the whole set of identified behaviors we found many agreements and no contradictions among the experts' opinions. Moreover, the behaviors identified in this study show an alignment with the vocal expression and nonverbal behavior items from the validated oral presentation rubrics presented in Schreiber et al. 2012. This overall agreement aligns with the purpose of the PT, which aims to support the development of basic skills. It does not aim to train professionals to learn and create their individual presentation style.

Technical nonverbal communication behaviors

This study was conducted in the context of improving the PT; therefore it is relevant to analyze the feasibility of implementing computerized mechanisms to recognize the identified behaviors. Regarding the set of postures identified postures, it is possible to recognize them using depth cameras such as the Microsoft Kinect sensor (Le, Nguyen & Nguyen, 2013; Xiao, Mengyin, Yi, & Ningyi, 2012). This type of cameras have also been used to recognize predefined gestures (Li, 2012; Patsadu, Nukoolkit, & Watanapa, 2012; Ren, Yuan, Meng, & Zhang, 2013).

Some of the gestures practices mentioned by the experts, e.g. "waving both arms above the shoulders", "crossing arms", etc. are predefined, i.e., can be described with clear spatial constraints, therefore techniques to recognize these predefined gestures can be used. However, practices such as "gestures bigger than usual", "delivered gestures", etc. are not predefined,

hence the amount of gestures that fall in this category is infinite and identifying them is still an open challenge.

"Vocalized gestures" can be identified through a multimodal. This approach requires input from microphones and depth cameras in order to identify whether a gesture is performed while the speaker is talking. By applying speech recognition techniques (Rabiner, & Juang, 1993; Graves, Mohamed, & Hinton, 2013) in combination with gesture recognition techniques it is possible to programmatically identify cases such as "Gestures for enumeration and sequences". This is by identifying predefined words such as "first", "second", etc., while the speaker is doing a corresponding gesture.

The automatic recognition of facial expressions is feasible as shown in the study of Bahreini, Nadolski and Westera, (2014). There are several techniques that can be used for eye tracking (Chennamma, & Yuan, 2013) that could be used for recognizing eye contact.

Regarding the voice behaviors identified by experts, there are existing techniques that can be used to recognize behaviors dealing with voice volume (Schneider et al. 2015b), voice pitch (Ghahremani, BabaAli, Povey, Riedhammer, Trmal, & Khudanpur, 2014) filler sounds (Prylipko, Egorow, Siegert, & Wendemuth, 2014), and voice emotion (Bahreini et al., 2014). Recognizing behaviors such as "talking out loud to yourself" instead of "speaking to the audience" and "stressing important words" remain currently an unsolved challenge.

The volume values captured by a microphone can be used to recognize pauses (Batrinca et al., 2013; Schneider et al., 2015b). Just by timely measuring the length of a pause, it is possible to differentiate between long and short pauses. However, experts did express not to know the length of short and long pauses, since they have always assessed these lengths intuitively without ever measuring them. Some solution to retrieve the lengths of shorts and long

pauses is by timely measuring the pauses in recorded presentations (Schneider, Börner, Van Rosmalen, & Specht 2015c). Automatically assessing the precise moment to deliver a short or long pause is currently an unsolved challenge. In order to correctly recognize these correct moments computers have to understand the content of the presentation, something that currently is not feasible.

Formative expert evaluation of the PT

The formative evaluation of the PT made us to reconsider the type of feedback given by the PT, and helped us to identify how tools such as the PT can enhance current practices for learning public speaking skills. Regarding the feedback of the PT, before this study the feedback of the PT was designed to teach learners how to present correctly. Nevertheless, experts recurrently remarked that there is no right way to do a presentation; therefore instead of being a corrective tool, experts suggested to design the PT as a tool to support learners with the development of awareness. To raise awareness some experts recommended the use of questions and warnings as feedback instead of corrective instructions.

In terms of the enhancement of current practices for learning public speaking skills, experts stated that students in public speaking would benefit by using a tool such as the PT. In the case of students following a public speaking course, teachers could give homework asking students to practice certain skills using the PT. In the case of seminars and intensive public speaking workshops, attendees could take the PT home, use it to prepare for future presentations and get reminded of the lessons learned during the intensive training sessions.

Finally, experts claimed that the PT cannot substitute a human tutor. We partially agree with this claim, because in a presentation the verbal and nonverbal communication are tightly coupled. Currently it is not feasible for computers to make sense of the content of a presentation

and analyze both forms of communication simultaneously. Thus in terms of the quality of assessment and feedback a tool such as the PT indeed is not able to compete against a qualified human tutor. However, we consider that tools such as the PT can still be used as tutors and support learners in the learning scenarios where human tutors are not available, e.g., in online courses and in informal learning situations.

Conclusions

The interviews conducted in this study allowed us to obtain crucial information for the improvement and further research on the PT and multimodal public speaking instructors in general. Even though generally speaking there is no “right” way to do a presentation, in this study we identified a wide agreement on good and ineffective nonverbal communication practices for public speaking. In total we identified a set of 61 ineffective practices and 70 good practices that can be taught to novice students in public speaking, which is the target group of the PT. Many of these practices can be recognized through the use of already existing computational techniques, making it possible to significantly expand the current rule-based model of assessment of the PT. Thus ensuring that practicing with PT will support learners in becoming better presenters.

The formative evaluation conducted in this study helped to shift the focus of the PT’s feedback. As experts suggested future versions of the PT should include a feedback designed to raise the learner’s awareness instead of just correcting them. This evaluation also pointed out how current practices for learning public speaking skills can be enhanced by tools such as the PT by presenting learners with opportunities to practice and rehearse the lessons learned in classrooms or seminars.

To continue with the improvement of the PT, the plan is to conduct a feasibility study regarding the implementation of the new assessment model of the PT together with the improvements suggested by the experts. The further step is to implement the identified improvements of the PT based on its feasibility and relevance.

As shown with Pericles funeral oration memorable presentations can lead to giant leaps for mankind. Becoming a great public speaker able to give memorable presentations is a complex task. Mastering the behaviors identified in this study is just one small step in becoming a great public speaker. Current technologies such as the PT present learners with the opportunities to become aware and master these behaviors.

Acknowledgements

The underlying research project is partly funded by the METALOGUE project. METALOGUE is a Seventh Framework Programme collaborative project funded by the European Commission, grant agreement number: 611073 (http://www.metalogue.eu).

References

Bahreini, K., Nadolski, R., & Westera, W. (2014). Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments*, 1-16.

Barmaki, R., & Hughes, C. E. (2015, November). Providing Real-time Feedback for Student Teachers in a Virtual Rehearsal Environment. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 531-537). ACM.

Batrinca, L., Stratou, G., Shapiro, A., Morency, L. P., & Scherer, S. (2013, August). Cicero-towards a multimodal virtual audience platform for public speaking training. In *Intelligent Virtual Agents* (pp. 116-128). Berlin, Heidelberg: Springer.

Bjerregaard, M., & Compton, E. (2011). Public speaking handbook. *Snow College, Supplement for Public Speaking*.

Campbell, K. S., Mothersbaugh, D. L., Brammer, C., & Taylor, T. (2001). Peer versus self assessment of oral business presentation performance. *Business Communication Quarterly*, *64*(3), 23-40.

Chan, V. (2011). Teaching oral communication in undergraduate science: Are we doing enough and doing it right? Journal Of Learning Design, *4*(3), 71-79.

Chennamma, H. R., & Yuan, X. (2013). A survey on eye-gaze tracking techniques. *arXiv preprint arXiv:1312.6410*.

Damian, I., Tan, C. S. S., Baur, T., Schöning, J., Luyten, K., & André, E. (2015, April). Augmenting social interactions: Realtime behavioural feedback using social signal processing techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 565-574). ACM.

Dermody, F., & Sutherland, A. (2015, November). A Multimodal System for Public Speaking with Real Time Feedback. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 369-370). ACM.

DeVito, J. A. (2014). *The essential elements of public speaking*. Allyn and Bacon.

Docheff, D. M. (1990). The feedback sandwich. *Journal of Physical Education, Recreation & Dance*, *61*(9), 17-18.

Gallo, C. (2014). *Talk Like TED: The 9 Public Speaking Secrets of the World's Top Minds*. Pan Macmillan.

Ghahremani, P., BabaAli, B., Povey, D., Riedhammer, K., Trmal, J., & Khudanpur, S. (2014, May). A pitch extraction algorithm tuned for automatic speech recognition. In proceedings of the *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* (pp. 2494-2498). IEEE.

Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In proceedings of the *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* (pp. 6645-6649). IEEE.

Hinton, J. S., & Kramer, M. W. (1998). The impact of self-directed videotape feedback on students' self-reported levels of communication competence and apprehension. *Communication Education*, *47*(2), 151-161.

Joint Quality Initiative. (2004). Shared 'Dublin' descriptors for short cycle, first cycle, second cycle and third cycle awards. Retrieved on: www.jointquality.nl/content/CompletesetDublinDescriptors.doc.

Le, T. L., Nguyen, M. Q., & Nguyen, T. T. M. (2013, January). Human posture recognition using human skeleton provided by Kinect. In *Computing, Management and*

Telecommunications (ComManTel), 2013 International Conference on (pp. 340-345). IEEE.

Li, Y. (2012, June). Hand gesture recognition using Kinect. In *Software Engineering and Service Science (ICSESS), 2012 IEEE 3rd International Conference on* (pp. 196-199). IEEE.

Morreale, S. P., & Pearson, J. C. (2008). Why communication education is important: The centrality of the discipline in the 21st century. *Communication Education*, *57*(2), 224-240.

O'Connor, H., & Gibson, N. (2003). A step-by-step guide to qualitative data analysis. *Pimatisiwin: A Journal of Indigenous and Aboriginal Community Health*, *1*(1), 63-90.

Parvis, L. F. (2001). The importance of communication and public-speaking skills. *Journal of Environmental Health*, *63*(9), 44-44.

Patsadu, O., Nukoolkit, C., & Watanapa, B. (2012, May). Human gesture recognition using Kinect camera. In proceedings of the *2012 International Joint Conference on Computer Science and Software Engineering (JCSSE),* (pp. 28-32). IEEE.

Prylipko, D., Egorow, O., Siegert, I., & Wendemuth, A. (2014). Application of Image Processing Methods to Filled Pauses Detection from Spontaneous Speech. *young*, *15*(3), 3.

Quianthy, R., & Hefferin, D. (1999). Speak Up! A College Competency Assessment Tool. *Popular Measurement*, *2*(1), 27-29.

Rabiner, L., & Juang, B. H. (1993). Fundamentals of speech recognition.

Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust part-based hand gesture recognition using kinect sensor. *Multimedia, IEEE Transactions on*,*15*(5), 1110-1120.

Rubin, H. J., & Rubin, I. S. (2011). *Qualitative interviewing: The art of hearing data*. Sage.

Schneider, J., Börner, D., Van Rosmalen, P., & Specht, M. (2015a). Augmenting the senses: a review on sensor-based learning support. *Sensors*, *15*(2), 4097-4133.

Schneider, J., Börner, D., Van Rosmalen, P., & Specht, M. (2015b). Presentation Trainer, your Public Speaking Multimodal Coach. In *Proceedings of the 2015 ACM International Conference on Multimodal Interaction* (pp. 539-546). ACM.

Schneider, J., Börner, D., Van Rosmalen, P., & Specht, M. (2015c). Stand Tall and Raise your Voice! A Study on the Presentation Trainer. In *Design for Teaching and Learning in a Networked World* (pp. 311-324). Springer International Publishing.

Schreiber, L. M., Paul, G. D., & Shibley, L. R. (2012). The development and test of the public speaking competence rubric. *Communication Education*, *61*(3), 205-233.

Smith, C. M., & Sodano, T. M. (2011). Integrating lecture capture as a teaching strategy to improve student presentation skills through self-assessment. *Active Learning in Higher Education*, *12*(3), 151-162.

Swan, M. (2012). Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, *1*(3), 217-253.

Thucydides. Pericles' funeral oration: Thucydides' history of the Peloponnesian War, Book II, xxxv-xlvi (Greek and English Edition).

Van Ginkel, S., Gulikers, J., Biemans, H., & Mulder, M. (2015). Towards a set of design principles for developing oral presentation competence: A synthesis of research in higher education. *Educational Research Review*, *14*, 62-80.

Xiao, Z., Mengyin, F., Yi, Y., & Ningyi, L. (2012, August). 3D human postures recognition using kinect. In proceedings of *2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC),* (Vol. 1, pp. 344-347). IEEE.

Appendix A

Ineffective Practices	# Of Experts mentioning the behavior	Good Practices	# Of Experts mentioning the behavior
I. Posture			
· Giving the back to the audience	7	· Feet between shoulder and waist width firmly on the ground	8
· Dancing	6	· Shoulders back and relaxed	8
· Hands in pockets	5	· Chin up	7
· Hands behind the back	5	· Facing the audience	7
· Hands touching hair	5	· Open posture	7
· Hands touching face	4	· Hands loose next to your body with palms facing the audience	2
· Crossing legs	4	· Neck back	2
· Hands grabbing and playing with something	4	· Hands together just above the belt, without interlacing	2
· Hiding yourself	4	· Posture where you feel at ease with yourself	2
· Fiddling with hands	4	· Point toes to audience	1
· Neck forward	3	· Arms relax, one hand grabbing the thumb of the opposite hand	1
· Crossed arms	3		
· Hunch	3		
· Closed posture	3		
· Standing with the bodyweight on one leg	2		
· One leg in front of the other	1		
II. Gesture			
· No gestures	7	· Gestures bigger than usual	5
· Waving both arms above the shoulders	4	· Delivered gestures	5
· Holding things	4	· Gestures for enumeration and sequences	5
· Touching face, hair, etc. without a specific purpose	4	· Gestures for emphasis	5
· Playing with notes	4	· Gestures to explain and paint the picture	5
· Holding hands without a specific purpose	3	· Make a gesture and return to your posture	5
· Crossing arms without a specific purpose	2	· Vocalize gestures	4
· Waving arms below the hips	1	· Slower gestures	4

III. Facial Expression				
· Blank face	9		· Smile	8
· Grinning like an idiot all the time	1		· Congruent with the content	7
· Lack of enthusiasm	1		· Show the emotion you want to transmit	4
IV. Eye Contact				
· No eye contact	8		· Screen the audience and give as much eye contact as possible	10
· Fixed eye contact	8			
· Reading	5			
· Give back to the audience	4			
· Facing screen	4			
V. Use of Stage				
· Stand behind the computer screen, desk or lectern	6		· Move with purpose	5
· Move constantly from one side to the other	4		· Stand in a place where you can be seen	4
VI. Voice				
· Talking out-loud to yourself	8		· Speak to the audience	7
· Be aware only of the content	8		· Breath from belly	4
· Filler sounds such as: hmm, ahm, etc.	5		· Stress important words	4
· Monotone voice	3		· Match the emotion with message you want to convey	4
· Speaking too fast	2		· A bit louder than usual	2
· Not loud enough	2		· A bit slower than usual	2
· Dropping volume end of the sentence	2		· Changes on voice volume	2
· High pitch	1		· Voice according to phases of the presentation	2
· Mumble	1		· Lower pitch Men	1
			· Higher pitch Women	1
			· Signaling new topic with higher pitch on first word	1
			· Make clear the end of each sentence	1
VII. Pauses				
· Not pausing	10		· Big pause after telling something important	6
· Hurrying up	7		· Big pause after asking a question	6
· No difference between small and big pause	4		· Big pause before starting next topic	5
			· Small pause after every sentence	2
			· Big pause letting people read the slide, before you talk about it	2
			· Every 3 to 5 sentences a big pause	1
			· Good timing	1
			· Longer pauses than usual	1
			· Chunking sentences and use small pauses between the chunks	1

VIII. Walking to the stage				
· Hurry to the stage	3		· Walk slow and confident while giving eye contact to the audience	3
· Shuffling	2			
· Negative self talk	1			
· Ignore that the Presentation already started	1			
IX. Settle in Time				
			· Take your time	10
			· Get grounded	5
			· Deep breaths	4
			· Claim territory	1
			· Stand closer to the audience	1
X. Introduction				
· Starting with high pitch	1		· A lot of eye contact (more than usual)	6
			· Lots of Pauses	4
			· Lots of voice variation (volume, pitch)	4
			· Speak loud	4
			· Theatrical	3
			· Open arms	3
			· Prepared start	2
			· Come close to the audience	2
			· Low pace	2
			· Enthusiasm (smile)	1
XI. Middle				
· Monotonous speech	7		· Change dynamics	8
· No stress on important words	2		· Less energy as in the beginning	3
· Not using the stage	2		· Look away when trying to remember something or after a rhetorical question, and then look back again	1
XII. Conclusion				
· Not having a full stop	6		· Big pause before giving it	8
· Not signify that is coming	5		· Slow and clear	6
· Ending with: "And that's it"	4		· Make yourself big (Open posture, arms extended)	3
· Losing energy	1		· Come closer to the audience	1
			· Keep breathing	1
XIII. Questions and Answers				
· Focus only on the person asking the question	3		· Acknowledge question to person who asked the question	4
· Pointing with a finger to the person asking	1		· Give answer to everybody in the audience	4
· Bad timing, not giving time for questions	1			